DOCUMENT RESUME

ED 047 340 CG 006 228

AUTHOR Silverman, Mitchell
TITLE An Analysis of a Comprehensive Evaluation Model for Guided Group Interaction Techniques with Juvenile Delinquents. Final Report.
INSTITUTION University of South Florida, Tampa. Institute III: Exceptional Children and Adults.
SPONS AGENCY Law Enforcement Assistance Administration (Dept. of Justice), Washington, D.C.
PUB DATE Dec 70
NOTE 83p.; Volume 2, Number 1

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Behavioral Objectives, Data Collection, *Delinquents, Group Therapy, Guidance Programs, *Interaction Process Analysis, Learning, Measurement Techniques, Scoring, *Techniques

ABSTRACT

Reported are the first phase activities of a longitudinal project designed to evaluate the effectiveness of Guided Group Interaction (GGI) technique as a meaningful approach in the field of corrections. The main findings relate to the establishment of reliability for the main components of the Revised Behavior Scores System developed to assess the basic processes taking place in settings using GGI as a therapeutic technique. This system is derived from the scoring categories, surscores, and scoring conventions develcped by Borgatta and Crowther (1965) in the Behavior Scores System. Subjects were members of two groups of male delinquents undergoing GGI at a correctional facility in Tallahassee in the summer of 197C. Results indicate that in general the system is sufficiently reliable to generate sound data, although data generated for the surscores and the single scoring convention used indicates that these components in the analysis system should not be used for gathering quantifiable data. The system should yield highly reliable data relative to the behavior occurring in therapy groups while providing measures of the learning processes accounting for such behavior. (Author/CJ)

ED047340

INSTITUTE
:EXCEPTIONAL CHILDREN AND ADULTS
UNIVERSITY OF SOUTH FLORIDA
TAMPA, FLORIDA

PAPERS and REPORTS

AN ANALYSIS OF A COMPREHENSIVE EVALUATION MODEL FOR GUIDED GROUP INTERACTION TECHNIQUES WITH JUVENILE DELINQUENTS

by

Mitchell Silverman

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

VOLUME 2, Number 1
1970

EC 006 228

One volume of Papers and Reports (P&R) of INSTITUTE III: Exceptional Children and Adults, is issued each academic year. The frequency of individual numbers will vary from volume to volume.

P&R serves the same function as "Technical Reports" or "Working Papers" of other institutions. This publication is an outlet for the ideas and research produced under the auspices of INSTITUTE III, and is a channel of communication between the personnel of INSTITUTE III and students and faculty of similiar interests at the University of South Florida and elsewhere.

Copies of individual numbers may be obtained free of charge by writing Papers and Reports, INSTITUTE III: Exceptional Children and Adults, University of South Florida, Tampa, Florida 33620; by phone (813: 974-2815); or by communication with individual authors.

INSTITUTE III STAFF:

James W. Barnard, Ph.D.
Acting Director

Basil L. Gaar, M.Ed.

William R. Blount, Ph.D.

Mitchell Silverman, Ph.D.

INSTITUTE III: Exceptional Children and Adults, is a unit of the Research and Development Center, USF, William H. Taft, Ph.D., Director.

FINAL REPORT

Grant No. NI 70-081

AN ANALYSIS OF A COMPREHENSIVE EVALUATION MODEL FOR GUIDED GROUP INTERACTION TECHNIQUES WITH JUVENILE DELINQUENTS

Mitchell Silverman

December, 1970

The research reported herein was performed pursuant to a contract with the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, United States Department of Justice.

INSTITUTE III: Exceptional Children and Adults
University of South Florida
Tampa, Florida 33620

FOREWORD

The purpose of the present report is focused on the presentation of results derived from a methodological study designed to develop an evaluation model for the analysis of Guided Group Interaction techniques as they are used in rehabilitating juvenile delinquents. While the results contained herein relate specifically to the feasibility of applying a given set of evaluative techniques and procedures within the confines of the group treatment program developed by the Division of Youth Services in the State of Florida, it is felt that the knowledge gained from this endeavor should have implications for assessing a wide range of diverse rehabilitative programs existing in the field of corrections today. Even though the present phase of research was not designed to demonstrate the efficacy of a given type of treatment over another, but rather the feasibility of doing evaluative research in a theraputic setting with complex interacting variables, it was hoped that the results of this pilot investigation would aid those investigators in the field whose primary objectives are to develop better means of helping the youthful offender.

The recognition of those individuals, without whose hard work and dedication the Phase I activities would never have been possible, is gratefully extended. The interest and effort of the research assistants assigned to this project cannot be over estimated. Special acknowledgement is given to Mr. Maurice E. Lucas who was instrumental in carrying out the objectives of the initial phase of research. In addition, the tireless work done by the assistants responsible for scoring videotapes; Mr. David R. Elman, Mr. Frederick S. Elson, Mrs. Jan Geraghty, and Mr. R. Hal Shigley receive special consideration. Recognition is also extended to Mr. John L. Morganthau for his work in the production of the videotapes at the treatment center. Two other assistants are also deserving of mention, Miss Bonnie Bang and Mr. John Fosselman who aided with the data preparation and analysis.

The interest, cooperation, and support provided by the dedicated staff of the Division of Youth Services was essential to the carrying out of this project. The encouragement of Mr. O. J. Keller, Director of the Division provided and continues to provide a motivational basis for the writer's work. The

consideration extended by Mr. Richard Rachin, Mr. Martin Cox, Mr. Guy Moore, and the entire staff of the group treatment program warrant very special thanks. Nor should the boys living at the Walter Scott Criswell House be overlooked for their job is hard enough and yet they participated in our research freely.

Next, my very special thanks is given to Mrs. Magdalene Deutsch, the project secretary. Her capable job of handling project records, and assistance in the preparation of manuscripts made the project director's role much less of a burden.

Last, appreciation is acknowledged to my colleagues in INSTITUTE III: Exceptional Children and Adults for acting as a sounding board to the many ideas researched in this project. To Drs. James W. Barnard, William R. Blount, and John N. Shadgett, and Mr. Basil Gaar, I give my thanks for providing an atmosphere conducive to the exploration of new ideas.

TABLE OF CONTENTS

	Page
Introduction.	1
Methods	17
Subjects	17
Apparatus.	19
Treatment.	20
Treatment Setting.	21
Procedures	22
Videotape Procedures	22
Videotape Scoring Procedures	25
Procedures for Time-Effort Analysis. . .	35
Results	37
Discussion.	55
Summary	70
References.	73

TABLES

Table No.	Titles	Page No.
I	Reliability Coefficients for Consistency Across Judges for Categories I through VI for Group A	38
II	Reliability Coefficients for Consistency Across Judges for Categories I through VI for Group B	39
III	Reliability Coefficients for Consistency Across Judges for Categories I through VI for Combined Groups A & B	40
IV	Reliability Coefficients for Consistency Across Judges in Using Positive Reinforcement for Categories I through VI for Group A	44
V	Reliability Coefficients for Consistency Across Judges in Using Positive Reinforcement for Categories I through VI for Group B	45
VI	Reliability Coefficients for Consistency Across Judges in Using Positive Reinforcement for Categories I through VI for Combined Groups A & B	46
VII	Reliability Coefficients for Consistency Acorss Judges in Using Negative Reinforcement for Categories I through VI for Group A	47
VIII	Reliability Coefficients for Consistency Across Judges in Using Negative Reinforcement for Categories I through VI for Group B	48

TABLES

Table No.	Title	Page No.
IX	Reliability Coefficients for Consistency Across Judges in Using Negative Reinforcement for Categories I through VI for Combined Groups A & B	49
X	Time Motion Data For Videotape Production Tasks	52
XI	Time Motion Analysis of Tasks Required for Videotape Scoring	54

FIGURES

Figure No.	Title	Page No.
1	Schemata of GGI Setting	23
2	Individual Subject Scoring Sheet	33

INTRODUCTION

Assessing the effectiveness of programs designed to rehabilitate delinquent youth must be considered one of the critical issues facing researchers in the field of corrections today. Even though great strides have been made in replacing traditional institutional settings in which rehabilitation primarily consisted of corporal punishment, little data can be found to document the validity of the new approaches over the old in terms of preparing the individual for productive roles in society upon his return to the community. Furthermore, even less data is available relative to the internal dynamics of the various treatment approaches being used which potentially are the critical variables in producing the behavior changes necessary if rehabilitation is to be successful. It is the primary purpose of the research presented in this paper to develop techniques and approaches that can be used to fill the data gap discussed above.

A review of the literature on the effectiveness of various treatment programs for delinquents revealed several trends representing different methods for collecting evaluative data

in this area. First, there were a great number of publications reporting the results of case studies or clinical observations as documentation for the efficacy of given theraputic techniques in bringing about changes in personality structure or behavior patterns. While the following publications represented a limited sample of this approach to evaluation, they were considered more-or-less characteristic of this trend of investigation. Glasser (1965) in his book on reality therapy used primarily case study information as a basis for the justification of his position. Glasser presented a number of detailed reports of his clinical observations relative to individual patients as they progressed through therapy along with supplemental information that had been gathered after therapy had ended (Glasser, 1965; chapter three). The study by Jacobs and Christ (1967) also represented the case study approach. These authors clinically assessed the effect of "structuring" and "limit setting" on facilitating group therapy with six delinquent boys. In a third study Peins (1967) discussed the case histories of ten mentally retarded delinquent boys who received client centered communication therapy. While

this study differed from the others in that the intervention focused on language development of the subjects, it still represents a good example of the clinical approach to evaluation.

The above authors provide a great deal of useful information relative to the implementation of new and different theraputic techniques. However, the data that they generate in defense of their positions is subject to criticism with respect to experimenter bias, sample selection procedures, sample size, lack of objective measures, quantifiability of the data, generality of the results, etc. The value of such research is huristic in nature, and investigators using other approaches should not over look the many fine ideas expressed in this aspect of the literature.

The second trend noted in the literature on research designed to evaluate rehabilitative programs for delinquents was characterized by those studies that assessed the effectiveness of a given theraputic technique in terms of personality adjustment as defined by scores on various types of personality tests. An example of this type of research was

found in the work of Sandhu (1966) who evaluated the effects of therapy specifically designed to strengthen the self concepts of delinquent boys in a reformatory. Thirty-three boys were pre- and post-tested on a series of tests containing the Thematic Apperception Test, the Socialization Scale of the California Psychological Inventory, a self-image questionaire, a personality word list, and a self written autobiography. On the basis of differences between the pre- and post-test scores on these measures the author concluded that the therapy succeeded in improving the self-images of the boys receiving the special treatment. It should be noted that an appropriate control group was not used in the study. Shore, Massimo, and Mack (1965) evaluated changes in the perception of interpersonal relationships in delinquent boys who had received special vocational oriented therapy. Experimental and control groups of ten subjects each were compared on the basis of responses made to cards selected from the Thematic Apperception Test, and the Emotions and Motivations Test developed by Seymore Epstein of the University of Massachusetts. The cards were administered after the experimental

group had received ten months of therapy. The authors concluded that the therapy did result in improvement of interpersonal relationships and that such changes were related to improvement in overt behavior and higher performance levels. A third study representing this type of evaluation was carried out by Persons (1966). This writer evaluated the psychological and behavioral changes occurring in delinquents after they had experienced intensive group and individual psychotherapy. Two groups of 41 subjects each were selected from a state reformatory and randomly assigned to either a therapy or control group. Both groups were tested twice, once before the experimental group received therapy and again at the conclusion of treatment. The measures used for the pre- and post-tests were the Delinquency Scale developed by Peterson, et al. (1959), the Taylor Manifest Anxiety Scale, and the Minnesota Multiphasic Personality Inventory. The author found that the group who received therapy showed significantly greater improvement in terms of institutional adjustment, better interpersonal relationships, school performance, and overt behavior.

The above studies have a number of advantages over those

reported earlier using the case history approach. First, methods they employ in collecting their data are much more amenable to systematic replication by other researchers than would be the case for clinical observation. Second, the measures in these studies, even in the case where projective tests are used, can be considered more objective in that the procedures for their administration have been well standardized. In those cases where accepted objective tests are given, this argument gains added strength in that not only are the procedures for giving and scoring the tests standardized, but the tests often are emperically validated on samples known to have the characteristics being studied. Third, the subjects selected in the above studies generally are more representative of the larger populations to which the authors wish to generalize, therefore, the results collected in these investigations have much broader implications. Last, this type of investigation frequently employs repeated measure designs which enables the investigator to attempt to assess the impact of the therapy in terms of producing change in the degree of pathology or adjustment.

In contrast the following criticisms can be made of the second trend of evaluative research noted in the literature. The initial criticism focuses on the fact that while such studies provide a large amount of information relative to personality and behavioral changes correlated with therapy, they still do not address themselves to the process level that accounts for such changes. Another major point is that the research discussed above does not present enough data points to accurately represent the temporal sequence of changes that do occur in treatment. Next, none of the above cited studies provide followup data on their subjects, therefore, the long term effects of a given treatment approach are not known. In addition, a number of methodological criticisms can be made about studies such as those reported above. Several of the investigations use tests that are not commonly cited in the literature without presenting reliability and validity data on the measures for the samples being evaluated. This makes the interpretation of their results somewhat tenuous. Another common error is that the specific procedures used in treatment are not elaborated on in enough depth, thereby making the

research much more difficult to replicate. Finally, in many cases sample selection appears to be determined more by the availability of given groups of subjects than by good sampling procedures. This oversight frequently limits the generality of the results from otherwise well designed projects.

A third trend in evaluative research was characterized by in depth research efforts directed at assessing total treatment programs using a combination of methods and measures. An example of this type of approach were the studies reported by Knight (1969, 1970). These studies provided an ongoing summary of a continuous research evaluation of the Marshall Program for delinquent boys directed under the auspices of the California Youth Authority. The Marshall Program consisted of an intensive short-term (90-day) residential treatment program in which the emphasis was on inducing boys to examine their deviant attitudes and behavior through the use of group interaction techniques in which all individuals of the treatment facility, staff and delinquents alike, participated. The goals of the program were designed to change the boys through inducing them to adopt new perspectives and behavior patterns that

would minimize their chances of exhibiting future delinquent behavior. The goals of the first study were to initiate exploratory research relative to differences existing between program graduates and a group of boys who had been transferred from the program because they were felt to be unsuitable for this type of treatment. In addition, Knight examined the parole performance of Marshall graduates with that exhibited by a control group of subjects who met all of the criteria for admission to the Marshall Program, but who underwent other types of treatment in other state institutions. These groups were compared in terms of overall group differences in parole violation rates, and group differences in violation rate according to relevant background variables.

The results indicated that sixteen percent of the boys assigned to the Marshall Program were eventually transferred to other programs with the justification being lack of program involvement, and resistive attitudes. When compared to the Marshall graduates the failures showed a disproportionate number of lone offenders, more extensive prior history of delinquency, younger age range, and more extensive history of

prior escapes. Their responses to items on the Jessness Inventory indicated a greater propensity toward future conflict with the law, and they were more likely than graduates to be alienated from peers, staff, and program. When compared to the control group matched on admission criteria, the Marshall graduates showed a slight tendency to do better on parole. However, when the selection bias was partially controlled, these differences disappeared. The Marshall subjects also tended to show a relatively stable parole outcome irrespective of when a given individual was admitted to the program. The author found that graduates admitted during the first half of the eighteen and one-half month study period were identical to those subjects admitted during the second half of the period with respect to parole violation records. Neither the graduates from the Marshall Program nor the comparison group could be differentiated from each other in terms of fluctuation in cumulative violation rates through the fifteen month period of parole. From these findings the author concluded that the graduates of the Marshall Program performed just as well in terms of parole violations as did

the comparison group of boys who were released after a longer more expensive institutional experience. It was also noted that in terms of background variables, a specific type of individual appeared to benefit most from the short-term program. This unique group consisted of those boys in the Marshall setting who were older and appeared to be more interpersonally responsive.

In the second report Knight (1970) discussed the findings of a more recent data collection period relative to a new Marshall Program release group. Similar to the study reported above, a special control group was selected from a pool of wards of the California Youth Authority who met the admission criteria for the Marshall Program. Again the goal was to investigate the differential parole records of the two groups with respect to specified background factors. The findings from the earlier study were replicated during the second evaluation period with the results indicating that those individuals most amenable to the short-term intensive confrontative experience characteristic of the Marshall Program possessed the greatest degree of social and personal maturity. In

contrast, the group who was most prone to failure in this type of program was characterized by poor interpersonal relationships, had repeated escape histories, and had little faith in the program. These and other findings have led the author to conclude that in an intensive confrontative group interaction program such as Marshall's where the individual was forced to probe himself and his peers the potential exists both for treatment growth and harm. Those individuals who possessed a relatively high degree of ego strength and social skill along with perceptivity were felt to be the best candidates for programs employing rehabilitative techniques similar to those used at the Marshall setting. Subjects who did not possess these traits were felt to be poor risks with respect to success in the program.

A number of positive comments can be made about the above summarized research. One of its strongest points is that it attempts to evaluate program effectiveness in terms of success after the delinquent is returned to the community. Since the goals of rehabilitative programs for delinquents generally stress the importance of post-treatment adjustment, the results

of this type of investigation yield highly important information. In addition, the data from these studies is collected in such a manner that comparisons are feasible between the Marshall Program and other treatment settings in terms of successful adjustment after the boy is released. This enables the author to make statements relative to the cost-effectiveness of one program in relation to the others. Next, there is a genuine attempt made to find out what type of individual most benefits from a confrontative short-term program in contrast to those individuals who are least likely to derive any positive impact from the program. Last, this research consists of a continuous evaluative effort which has the advantage over studies done in isolation in that the results of one segment naturally provide the clues for successive phases of investigation.

While the above research in some respects is the most systematic of all the studies cited, it still contains a number of data gaps that have to be eliminated before the total picture is completed. Even though a great deal of effort is spent on collecting post-treatment data on the subjects behavior, this

information is not related to the ongoing behavior taking place during the group interaction sessions. Because of the differential subject selection procedures employed in selecting candidates for the Marshall Program, it is hard to differentiate between success attributable to treatment and success attributable to admission requirements of the Marshall Program. Similarily, the transfer of boys found to be unsuitable for the program raises serious questions about any claims made about the Marshall Program's effectiveness. It could be argued that only those boys are kept in the program who have those traits which are known to guarantee success after a given individual is released. If the population of delinquents at the Marshall Program did truly represent the broad spectrum of boys that ultimately come into contact with the law, the results from these studies would have had a much greater impact in providing some of the answers needed in the field of corrections today.

From the above discussion it is apparent that many variables still have to be investigated before we have effectively evaluated the potential of the increasing number of rehabilitative programs for delinquent youth. New evaluative research

efforts must be designed in a manner that enables the researcher to isolate the basic processes occurring during treatment so that their effects can be measured. Attempts also must be made to study the relationships between these processes and changes in personality and behavior that occur through out the treatment period. In addition, the predictive relationships between the events taking place during rehabilitation and subsequent adjustment after the individual is released from a given program and is living back in the community must be carefully investigated. If these goals are accomplished, researchers in the area of corrections will be much closer to the answers they have been seeking relative to the utility of different approaches to rehabilitation.

The purpose of the present research relates to the initiation of a comprehensive research effort designed to accomplish the tasks elaborated on above. The critical problem facing the researcher trying to deal with these issues relates to the development of a methodology that can be used to analyze the basic processes taking place during ongoing theraputic interventions. Measuring changes in personality or behavior taking

place during the theraputic period, while important, is not the same as analyzing the basic processes leading to such changes. The procedures and results discussed in the following sections of this paper constitute the first phase activities of a longitudinal research project designed to evaluate the effectiveness of Guided Group Interaction techniques as a meaningful approach in the field of corrections. The phase one activities centered on a feasibility study designed to analyze the utility of a complex interaction analysis system developed to measure the group processes that lead to behavioral changes occurring during Guided Group Interaction sessions. The specific objectives of the first phase of research can be given as follows:

A. Qualitative analysis of the videotape procedure in order to assess the feasibility of recording an accurate representation of the behavior occurring during GGI.

B. Comprehensive evaluation of the procedures designed for the analysis of the ongoing behavior recorded on the videotapes.

C. Evaluation of the time required in terms of personnel hours for the scoring of the videotapes.

D. Analysis and revision of the training procedures for both the videotape and scoring assistants.

E. Field testing of other behavioral measures to be used

in the Criswell House Setting.

F. Evaluation and revision of the initial data analysis procedures.

METHODS

Subjects

The subjects selected for the feasibility study consisted of all the members of two groups of male delinquents undergoing Guided Group Interaction (GGI) at the Walter Scott Criswell House in Tallahassee, Florida during the time period starting on July 1, 1970, and ending on September 30, 1970. In addition, the group leaders responsible for directing the GGI sessions were also included in the sample in that the interaction analysis system being field tested treats the therapist as a functioning member of the group structure. Because the purpose of the research being conducted during the first phase of operations was to assess the adequacy of the interaction analysis system, and not the efficacy of the treatment program, the random selection procedures that are to be used in subsequent phases of the research were not employed.

The first group (Group A) consisted of nine boys who had been adjudicated delinquent in the courts of Florida. These subjects ranged in age from 15 to 18 years, and their records indicated that the offenses for which they were convicted varied from incorrigibility to aggravated assault. The racial breakdown in the group was eight white subjects and one black subject. Six of the boys came from families that were intact, two from families where the parents were separated, and one where the father was deceased. The socio-economic level of the families from which these boys came ranged from a low of under 5,000 dollars annually to a high of well over 15,000 dollars in annual income. Similarly, all of the boys in the second group (Group B) were adjudicated delinquents who had committed offenses of varying severity. The age range for the seven boys in this group ranged from 15 to 16 years, with all of the boys in Group B being white. Five of the subjects in the second group came from intact families, one from a home where the father was deceased, and no information was available on the seventh subject in this group. The annual income of Group B ranged from a low of from 5,000-7,000 dollars to a high of over 15,000 dollars

per year.

The group leaders included in the sample were employed staff members of the Division of Youth Services in the State of Florida, who had received extensive specialized training in conducting GGI groups under the direction of Mr. Richard Rachin, Director of the Group Treatment Program for the division. Both of these individuals were male and were between 25 and 30 years of age. In addition to the specialized training received from the division, each of the GGI leaders were enrolled as students at Florida State University in Tallahassee working towards degrees in the field of criminology.

Apparatus

One complete videotape studio unit was used during the first phase of operations for the taping of the ongoing group processes occurring during GGI at the Criswell House facility. This unit consisted of a videotape recording deck designed to use one-half inch tape, two cameras equipped with external synchronization and wide angle lens, a synchronization generator, two omni microphones, a microphone mixer, and supportive equipment (e.g., cable, extension cords, stands, etc.).

In addition, two units were used to score the videotapes after the GGI groups were recorded. Each unit consisted of a videotape deck designed to use one-half inch tape, one 19 inch monitor, earphones, and supportive equipment (e.g., cable, extension cords, remote controls, etc.).

Treatment

A number of specific parameters were derived that best define the important processes taking place during the GGI sessions at the Criswell House. Basic to all activities carried out within the context of the groups was the emphasis placed on the variable of learning and/or re-learning the appropriate behavior responses to the demands placed upon the individual group member by society. Central to the issue of learning was the role which the group plays in directing the individual member's behavior through the interaction processes taking place during the ongoing therapy. One of the main processes used to accomplish this task was confrontative techniques, whereby the individual was forced to discriminate whether or not the responses he made to everyday situations were appropriate. Related to the confrontative nature of GGI was the strong emphasis placed on

the relationship between the group member's responses in terms of the reality of the world in which he comes from, lives in, and must return to after he returns to the community. Another important factor that had to be considered was the supportive relationship established between the members of the GGI group. This relationship gains added importance in that it provides the basis for the effectiveness of peer pressure in bringing about the behavioral changes in members of the group through the use of social reinforcement.

Treatment Setting

The setting used to tape the GGI groups for the present study consisted of a large room which, in addition to being used for therapy, was used for recreational and study activities. This room was selected for taping in that it provided several unique advantages. First, there was sufficient room available to adequately maneuver the cameras so that the possible recordings could be made of the interaction taking place during the GGI meetings. Second, the room had sufficient lighting so that no additional lights were necessary during taping. In that lights generate a great deal of heat, this helped minimize the

effects of extraneous stimuli interfering with the GGI. Third, this room had a projection booth directly adjacent to it which enabled the videotape operator to record the sessions with only the cameras being present in the room. A complete representation of the treatment setting including seating arrangements and camera placements is presented in Figure 1.

Procedures

Because of the complexity of the methods used in the feasibility study, the procedures for the different components will be discussed separately in order to facilitate clarity. In those cases where the procedures overlap across components, every effort will be made to make these relationships explicit.

Videotape Procedures: Upon initiation of the Phase I activities, the videotape equipment was assembled at the Criswell House setting and operationally tested. During the equipment assembly period, the videotape research assistant received specialized training in the use of the equipment from a videotape expert from the University of South Florida. This training consisted of teaching the assistant production techniques related

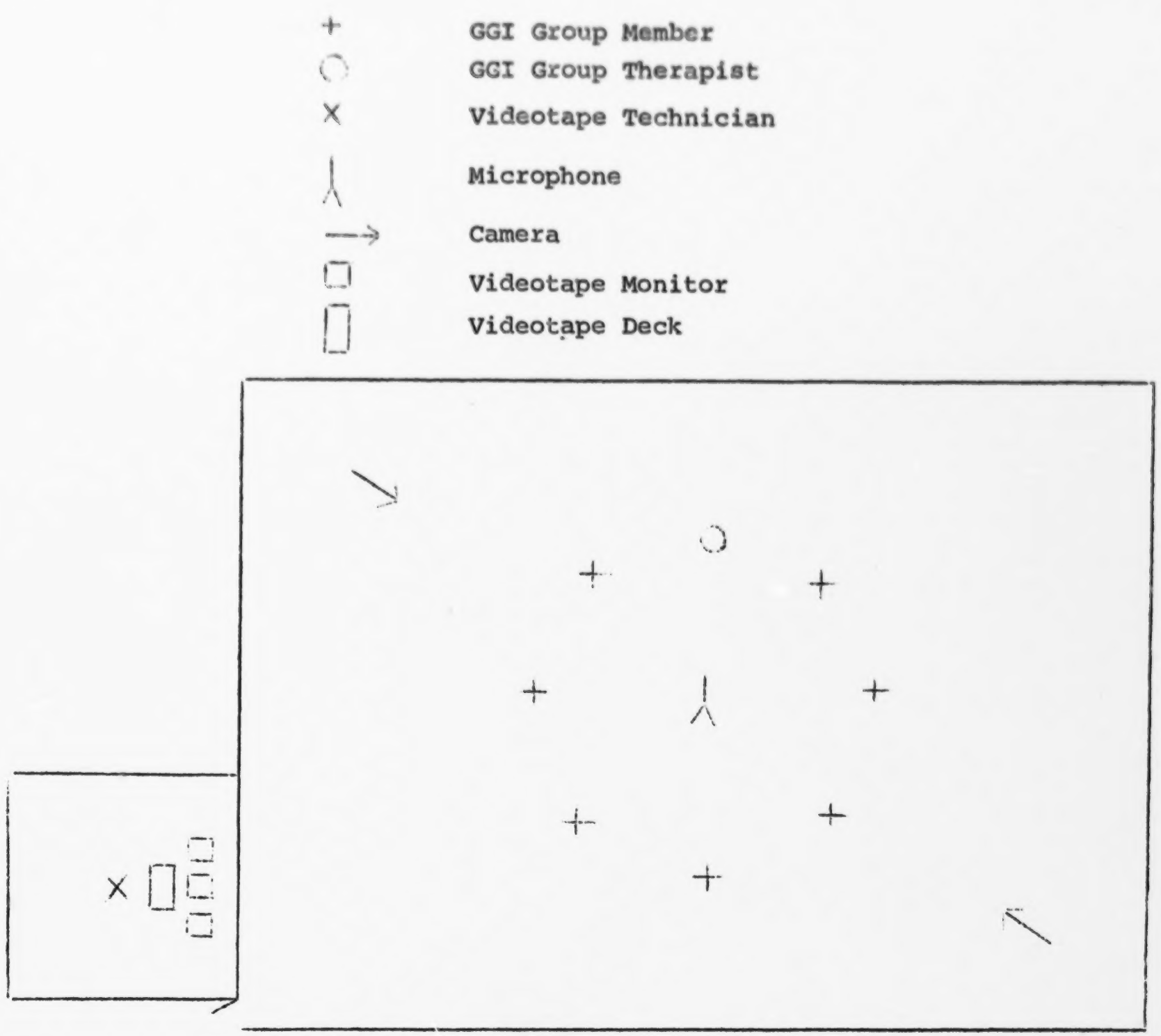

Figure I: Schemata of GGI Setting.

to camera placement, camera focussing, lighting, special effects, sound production, etc. In addition, the videotape assistant was trained in light maintenance of the equipment in an attempt to minimize the down-time required for equipment repair.

The research assistants responsible for scoring the tapes made at Criswell House also received special training in the use of the equipment used for tape playback. This training consisted of learning how to tune the units, equipment assembly, and light maintenance. Before any of the assistants were allowed to use any of the equipment they had to receive training from the videotape specialist.

Upon completion of the training, installation and testing of equipment, one complete therapy session per week for each of the GGI groups was taped for an eight week period. Each session required two complete tapes so that a total of 36 tapes were made during the eight weeks that data was being collected. After the tapes for a given session were produced, the videotape assistant made a diagram of the seating arrangement noting the position and names of the members comprising the group. The tape was then labeled with respect to group designation, session

number, and the date that the tape was made. In addition, the videotape assistant was required to maintain a log related to any significant events that occurred during the sessions he was taping. These records subsequently yielded a great deal of useful information relative to production difficulties, corrections in procedures, and suggestions for revising the videotape methodology. After the tapes were coded they were sent to the scoring laboratory located at INSTITUTE III: Exceptional Children and Adults at the University of South Florida.

Videotape Scoring Procedures: After the tapes were sent to the INSTITUTE they were scored in terms of the interaction processes taking place in the GGI groups by means of a revised form of the Behavior Scores System originally developed by Borgatta and Crowther (1965). This system contained three primary scoring categories in which an attempt was made to classify the following forms of responses: 1) Assertive Actions, 2) Withdrawal, and 3) Supportive Actions. The three classifications were further divided into specific behavioral subcategories and were defined as follows (Borgatta and Crowther, 1965; p. 50):

Assertive Actions

1. Neutral Assertions or Communications: These responses are basically continuations, explanations, expositions, and other forms of communication that add to the amount of talking, activity, and maintenance of a prominent and/or visible position in the communication process of the group. Responses scored in this category, in a sense, may be defined as "filler" on the part of the visible initiator.

2. Assertions or Dominant Acts: These are defined as acts in which the individual takes the prominent position, initiating conversations, and altering the pattern of discussion being carried out in the group.

3. Antagonistic Acts: These acts that are primarily associated with the rejection of others, and/or the rejection of others through the rejection of positions that others take. Self-assertive and ego-defensive acts tend to reflect the rejection of others, therefore, in those areas where a response has the connotation of self-assertiveness and/or ego-defensiveness, the appropriate category is Antagonistic Acts.

Withdrawal

4. Withdrawal Acts: Responses scored under this category are defined as leaving the field, failing to respond when the situation demands it, or unsuccessful attempts to enter the conversation or discussion. Withdrawal does not necessarily require an increase in tension on the part of the individual, although this is often the case.

Supportive Actions

5. Supportive Acts: Supportive acts may be defined as agreements in which the implication is "all right, go ahead and continue," or "I hear you".

6. Assertive Supportive Acts: These responses are differentiated from supportive actions in that the responses imply direct agreement with another group member's statement, thus raising the status of the other. In addition, other forms of status raising are scored within this category. Assertive supportive responses go beyond mere responsiveness in the sense that initiative is taken by the individual in support of the other, or his position.

In terms of evaluating the behavior taking place in the GGI groups, every action taking place during a given session was scored in one of the six categories summarized above. Added precision in scoring responses was achieved by using the qualitative surscores recommended in the original scoring system. The Group Oriented and Emotional Quality Surscores were defined as follows (Borgatta and Crowther, 1965; p. 50.):

Group Oriented Surscores

a. Task Determining Acts: These are defined as acts that draw attention to the task of the group, return the group to the task considerations, or move the group on the task to a further concern.

b. Group Maintaining Acts: In contrast to Task Determining Acts, responses scored under this category are directed at drawing the group together, raising unity, breaking deadlocks, etc.

Emotional Quality of Actions Surscores

c. Tension Displayed: A response is scored under this category if there is a noticeable increase in tension as evidenced by displays of nervousness, anxiety, pressured behavior, etc.

d. Unpredictable Behavior: A response scored under this classification is defined as representing an over-reaction, over intense reaction, an emotional display, a non-conventional reaction, or autistic and unrelated action implying lack of contact with the inter-action taking place in the group.

In addition to the surscores discussed above, three conventional response scoring categories were recommended in the original system. However, after using the system in the training sessions it was felt that two of the response conventions were redundant in that these classes of behavior were already covered under the surscore classifications. The two conventions that were dropped were the Convention for Withdrawal Under Tension, and the Convention for Withdrawal Under Obvious Hostility. The one remaining convention was essentially kept intact using the same definition given by Borgatta and Crowther

(1965, p. 50.) with only the designation being changed for clarity. The definition for this response convention can be given as follows:

Conventional Response Scoring

> Z. Convention for "...ah..." and False Starts Continued Successfully, are not Interpretable as Withdrawals. This category is not to be used when the above behavior is a matter of language pattern on the part of the individual and does not appear to be a nervous response.

The scoring system discussed above was selected primarily because it provided a set of well defined categories into which a broad range of behavioral responses could be classified enabling the researcher to keep track of the relative frequency of response rate in terms of objectively defined behaviors. However, since the primary goal of the present research endeavor was to develop a system that was not only sensitive to the phenomenon occurring within the context of a group treatment setting, but also to assess such phenomenon in terms of the processes that facilitated their occurrence, the original system required additional modification. In order to accomplish this task, a basic assumption had to be made about the essential processes taking place during the GGI group meetings. This

assumption was derived from the increasing evidence provided by contemporary researchers in the area of learning that the majority of behavior patterns can be modified through the systematic use of reinforcement contingencies (Skinner, 1953, 1957; Ferster and Skinner, 1957; Ferster, 1953; Sidman, 1960; among others). If these findings can be considered valid then the assumption can be made that the behavior changes taking place during ongoing GGI resulted from the reinforcement contingencies used by the group leader and members of the group in response to the behavior in question. In order to isolate these reinforcement contingencies it was necessary to make additional modifications to the Behavior Scores System. The following modes of reinforcement were defined in order to account for the processes of learning taking place in the group settings.

> Positive Reinforcement: Positive reinforcement can be defined as any reaction by one or more members of the group that positively supports a given response made by an individual member of the group. This support may take the form of agreement with another's action, personally rewarding comment, etc.
>
> Negative Reinforcement: Negative reinforcement is defined as any reaction by one or more members of the group that rejects a given response made by an individual member

of the group. Negative reinforcement may take the form of disagreement with another's position, refuting the logic of another's response, personal rebuke, etc.

Non-Reinforcement: This contingency accounts for that case where the response of a given individual member of the group is not followed by any reaction on the part of other group members.

By using the reinforcement contingencies discussed above it was hoped that the analysis of responses made by an individual in the group in terms of the reactions to such responses by other members of the group would be possible. In the system outlined above, a basic relationship exists between the response modes and the reinforcement contingencies as they were used in the present investigation. This relationship centers on the fact that a given behavior pattern exhibited by one of the subjects can be scored both in terms of one of the response modes and one of the reinforcement contingencies. By analyzing the total interaction processes occurring in GGI in terms of both response and reinforcement the relationship between behavior occurring in the group and the processes shaping such behavior can be accounted for. A complete diagram of the design of the

revised scoring system can be found in Figure 2.

Before the scoring system presented above was used on data included in the feasibility study, each of the four scorers used for this task were required to undergo an extensive training period. The training program lasted for approximately four weeks and consisted of a number of different tasks designed to familiarize the scorers with all aspects of this complex model of interaction analysis. The training consisted of learning the definitions for all of the response mode and surscore categories along with a comprehensive review of reinforcement therapy. In that all of the scorers selected for the project were graduate students in the field of psychology they were already grounded in the basic principles of learning theory which made the training task much easier.

Upon completion of the training period all of the scorers began to analyze the tapes made at the Criswell House setting. Each tape was scored by all four of the scorers during this phase of research in order that reliability estimates could be made across scorers in using the Revised Behavior Scores System for both of the GGI groups. None of the scorers were allowed

	RESPONSE MODE	REINFORCEMENT MODE		
ASSERTIVE ACTIONS	frequency	positive	negative	none
1. Neutral Communication				
2. Assertive or Dominant Acts				
3. Antagonistic Acts				
WITHDRAWAL				
4. Withdrawal				
SUPPORTIVE ACTIONS				
5. Supportive Acts				
6. Assertive Supportive Acts				

a. Task determining acts

b. Group maintaining acts

c. Tension displayed

d. Unpredictable behavior

z. Convention for false starts

Figure 2. Individual Subject Scoring Sheet

to communicate with each other until a given tape was scored by all of the assistants and the data was entered and being processed. Each assistant was required to keep a log of the significant events and problems that occurred during the analysis of the tapes and from this information a number of important clarifications were made with regards to the basic definitions used in the system. At the end of each week a general meeting was held for all of the scorers at which time the issues recorded in the log books and any other questions regarding the scoring system were discussed. On several occasions the videotape assistant was included in these meetings when the scorers had some pertinent information and feedback regarding the production of the tapes.

All of the data collected in the above manner was summarized by session for each individual subject and by groups and entered onto master scoring sheets. From these sheets the data was keypunched and submitted for computer processing. Because of the logistics involved in handling this amount of data, it was necessary to carry out the procedures of summarizing immediately upon completion of the scoring by a given scorer.

Procedures For Time-Effort Analysis: In addition to the basic evaluation related to the use of the Revised Behavior Scores System, one of the primary emphases of the present research was to assess the utility and economy of the total set of procedures to be used in the remaining phases of research. This facet of the feasibility study was accomplished through the analysis of the effort required to carry out the major tasks included in the total evaluation of the GGI program. In that none of the procedures involved in the research design required great amounts of physical effort, the primary concerns in this aspect of the investigation related to the analysis of the amount of time and mental stamina required to complete a given task.

All components of the videotape production operation were subjected to time and effort assessment. The main tasks required in the production of the tapes at the GGI setting were the assembly of the cameras and microphones in the meeting room, adjustment of the camera with respect to focus, brightness, and position, adjustment of microphones for optimal sound production, preparing the tape for recording, checking out the entire system

prior to taping, fine tuning the entire system, changing tapes in the middle of the session, disassembling the equipment at the conclusion of the session, and coding tapes for the scorers. In that, with the exception of the cameras and microphones, all of the other equipment was permanently assembled, the task of putting together the other components was not evaluated. The log that the assistant kept was used as the source of information related to the mental requirements of this position.

Similarly, the videotape scoring procedures were also evaluated in terms of time and effort. The tasks comprising this aspect of the research related to the time required to review the tape completely prior to scoring, the time required to actually score one of the 45 minute tapes, the time required to summarize the data after scoring, and the time required to enter the data onto the master scoring sheets. The research assistants doing the scoring were required to note comments about any difficulty in scoring resulting from fatigue, boredom, and other factors.

In that a number of other measures will be used in the subsequent phases of research, trial runs were made with these

indices in order to evaluate the amount of time required to administer these measures. These measures included an alienation scale developed by Spilka (1970), The Wechsler Tests for intelligence for both adults and children, and a behavior rating scale completed by both selected group members and professional staff.

RESULTS

All of the data collected with respect to the Revised Behavior Scores System was used to establish reliability estimates for all of the components comprising this complex method for assessing interaction processes. Reliabilities for the response mode categories, reinforcement contingencies, surscores, and scoring convention were derived using the Hoyt Procedures for computing the Kuder-Richardson$_{20}$ Reliability Coefficient as summarized in Winer (1962, pp. 124-132). The coefficients computed for the six basic response categories across the four scorers for all seven of the sessions recorded can be found in Tables I, II, and III.

Table I contains the reliabilities derived from the data collected on the behavior taking place during the GGI meetings

TABLE I

Kuder-Richardson$_{20}$ Reliability Coefficients for Frequency of Response Across Four Judges Categories I through VI for Group A Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session I	0.98	0.96	0.74	0.64	0.78	0.78
II	0.95	0.95	0.98	0.86	0.63	0.74
III	0.98	0.99	0.90	0.91	0.87	0.90
IV	0.98	0.99	0.97	0.75	0.46	0.94
V	0.95	0.99	0.94	0.75	0.75	0.84
VI	0.99	0.96	0.66	0.74	0.34	0.67
VII	0.99	0.97	0.97	0.78	0.86	0.98

*The K-R$_{20}$ reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

TABLE II

Kuder-Richardson$_{20}$ Reliability Coefficients
for Frequency of Response Across Four Judges
Categories I through VI for Group B
Sessions I through VII

	Response Category					
Session	1	2	3	4	5	6
I	0.97	0.96	0.71	0.53	0.83	--0.02
II	0.98	0.96	0.97	0.91	0.96	0.65
III	0.95	0.95	0.88	0.96	0.86	0.63
IV	0.95	0.98	0.82	0.80	0.65	0.95
V	0.99	0.99	0.92	0.91	0.89	0.87
VI	0.99	0.54	0.97	0.94	0.70	0.92
VII	0.98	0.98	0.89	0.88	0.86	0.92

*The K-R$_{20}$ reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

TABLE III

Kuder-Richardson$_{20}$ Reliability Coefficients
for Frequency of Response Across Four Judges
Categories I through VI
Combined Groups A & B
Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session						
I	0.97	0.96	0.70	0.59	0.91	0.36
II	0.97	0.96	0.97	0.87	0.87	0.71
III	0.96	0.97	0.88	0.96	0.87	0.80
IV	0.96	0.98	0.90	0.81	0.55	0.94
V	0.98	0.99	0.96	0.85	0.89	0.86
VI	0.99	0.64	0.94	0.94	0.59	0.83
VII	0.99	0.98	0.93	0.86	0.85	0.97

*The K-R$_{20}$ reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

for Group A. One of the important findings was the high degree of consistency in reliabilities across sessions with respect to the way that the scorers interpreted the behavior they were analyzing. This was particularly the case for the reliabilities derived for the first two categories, Neutral Communication and Assertive or Dominant Acts, which showed a minimum of variation from session to session. Several inconsistencies were noted with respect to marked changes in the scoring reliability from one session to another (Category V for Session IV; and Categories III, V, and VI for Session VI). However, by the last session the reliability coefficients ranged from .78 for the IV category to .99 for the Neutral Communication category.

The reliability coefficients derived from the data collected on the behavior taking place during the GGI meetings for Group B can be found in Table II. The results for the second group were remarkably similar to those found for Group A, in that a high degree of consistency in reliability was noted with respect to the way the scorers analyzed behavior occurring in the group meetings from one session to the next. Again the most consistent coefficients across sessions were found for the

first two categories. The exceptions in consistency of reliability for the Group B data were categories IV and VI during the first session, and category II for the sixth session. By the last week of scoring the reliabilities ranged from .86 for the Supportive Acts category to .98 for both the Neutral Communication and Assertive or Dominant Acts categories.

After the reliabilities for Groups A and B were computed separately, the data for both groups were combined for further analysis. The reliability coefficients derived from this data for the six basic response categories for all sessions was summarized in Table III. As was expected the pooled data yielded the same trends that were noted for the individual group data summarized in the first two tables. The reliabilities for the six classes of behavior were very consistent from one session to the next, with the first two categories showing the least session-to-session variability. With respect to this analysis the exceptions were not as dramatic as those found for the individual group data, however, the inconsistencies could still be noted (i.e., Category VI, Session I; Category II, Session VI). These results show the effects of increasing the frequency of responses

on the procedures used to compute the reliability coefficients.

Once the data was evaluated with respect to the reliability of the six response mode categories, coefficients were then derived for the number of positive reinforcements per-category of response as analyzed by the scorers. The results of these analyses for the data derived from the two GGI groups for all the sessions recorded were summarized in Tables IV, V, and VI. As the results presented in these tables indicated, the reliabilities for the assessment of the number of reinforcements for a given category were much more variable and inconsistent across time. Only two of the six response categories received sufficent numbers of positive reinforcement to even warrant computing the reliability coefficients (Categories I, Neutral Communication; and II, Assertive or Dominant Acts). The number of positive reinforcements scored for the remaining categories were so few, and erratic that analysis of them was felt to be invalid.

Similarly, reliability coefficients were computed for the number of negative reinforcements scored per-response category, with the data being presented in Tables VII, VIII and IX.

TABLE IV

Kuder-Richardson$_{20}$ Reliability Coefficients for Frequency of Positive Reinforcement per Frequency of Responses Across Four Judges Categories I through VI Group A Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session						
I	--.	0.31	--	--	--	--
II	0.57	0.93	--	--	--	--
III	0.76	0.83	--	--	--	--
IV	0.74	0.83	--	--	--	--
V	0.69	0.52	--	--	--	--
VI	0.72	1.00	--	--	--	--
VII	0.33	0.89	--	--	--	--

*The K-F_{20} reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

TABLE V

Kuder-Richardson$_{20}$ Reliability Coefficients for Frequency of Positive Reinforcement per Frequency of Responses Across Four Judges Categories I through VI Group B Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session I	0.86	--	--	--	--	--
II	0.79	0.59	0.64	--	--	--
III	0.87	0.73	0.53	--	--	--
IV	0.52	0.59	0.69	--	--	--
V	0.65	0.85	--	--	--	--
VI	0.75	0.92	--	--	--	--
VII	--	0.64	0.48	--	--	--

*K-R_{20} reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

TABLE VI

Kuder-Richardson$_{20}$ Reliability Coefficients for Frequency of Positive Reinforcement per Frequency of Responses Across Four Judges Categories I through VI Combined Groups A & B Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session I	0.86	0.55	--	--	--	--
II	0.76	0.79	0.66	--	--	--
III	0.81	0.80	0.44	--	--	--
IV	0.66	0.72	0.51	--	--	--
V	0.81	0.78	--	--	--	--
VI	0.73	0.84	--	--	--	--
VII	0.52	0.79	0.33	--	--	--

*The K-R$_{20}$ reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

56

TABLE VII

Kuder-Richardson$_{20}$ Reliability Coefficients for Frequency of Negative Reinforcement per Frequency of Response Across Four Judges Categories I through VI Group A Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session						
I	0.70	0.65	-0.33	--0.77	--	--
II	0.94	0.96	0.90	--	--	--
III	0.74	0.91	0.64	0.48	--	--
IV	0.85	0.96	0.77	--	--	--
V	0.78	0.99	0.89	--	--	0.66
VI	0.88	0.78	--	--	--	--
VII	0.81	0.93	0.75	0.85	--	--

*The K-R$_{20}$ reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

TABLE VIII

Kuder-Richardson$_{20}$ Reliability Coefficients for Frequency of Negative Reinforcement per Frequency of Response Across Four Judges Categories I through VI Group B Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session						
I	0.31	0.95	-1.24	--	--	--
II	0.88	0.95	0.69	0.52	--	--
III	0.82	0.87	0.62	0.74	--	--
IV	0.80	0.72	--	0.39	--	--
V	0.54	0.85	--	--	--	--
VI	0.86	0.97	0.44	--	--	--
VII	0.77	0.87	0.52	0.86	--	--

*The K-R$_{20}$ reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

TABLE IX

Kuder-Richardson $_{20}$ Reliability Coefficients for Frequency of Negative Reinforcement per Frequency of Response Across Four Judges
Categories I through VI
Combined Groups A & B
Sessions I through VII

	Response Category					
	1	2	3	4	5	6
Session						
I	0.63	0.72	-0.60	-0.33	--	--
II	0.92	0.72	0.84	--	--	--
III	0.80	0.88	0.62	0.75	--	--
IV	0.82	0.83	0.55	0.35	--	--
V	0.77	0.99	0.90	--	--	0.66
VI	0.86	0.95	0.49	--	--	--
VII	0.77	0.87	0.69	0.85	--	--

*The $K\text{-}R_{20}$ reliability estimates were computed using the Hoyt procedures given in Winer (124-132).

The results presented in these tables indicated that the assessment of the number of negative reinforcements used by the scorers for a given response category was more consistent and less variable than was the case for the positive reinforcement contingency. In the case of the negative reinforcement contingency there were a sufficient number of responses to compute the reliability coefficients for at least three categories, and possibly a fourth (I, Neutral Communication; II, Assertive or Dominant Acts; III, Antagonistic Acts; and IV, Withdrawal Acts). As was the case with reliabilities discussed above for the six basic response categories, the first two classifications of response were the most consistent in terms of scoring the negative reinforcement contingency. The least consistent set of reliabilities were found for the fourth category of behavior ranging from -.33 for the first session to .85 for the last scoring period as indicated by the data presented in Table IX for the Combined A and B groups.

Assessment of the reliability of using the qualitative sur-scores and the scoring convention for false starts in conjunction with the six response modes in evaluating the behavior

taking place during the GGI sessions was also attempted. It was found that these scoring modes were used so infrequently by the scorers that it was impossible to compute reliabilities for them in relation to a given category of response. However, the scorers did feel that the surscores and false start convention were useful with respect to helping define a sequence of behavior when used in the context of the dynamics taking place during GGI. Further discussion of the descriptive function of these variables will be made in the following section.

In addition to the data analyzed for the reliability estimates discussed above, data was also collected relative to the time and effort involved in implementing the scoring system used in the present research. The data compiled relative to the production of the tapes at the Criswell House setting can be found in Table X. As noted in the table the time required to carry out all of the tasks relative to the videotape production operation was evaluated for three of the seven sessions being recorded (Sessions I, III, and VI). The average time for each task over the three sessions was 34 minutes for equipment assembly prior to taping, 13 minutes to check out the system

TABLE X

Time Motion Data For
Videotape Production Tasks

	Time Per Session			Average Time
Task	I	III	VI	
Equipment Assembly	40 min.	33 min.	34 min.	34 min.
System Inspection	10 min.	15 min.	15 min.	13 min.
Time to change tape during session	4 min.	3 min.	3 min.	3 min.
Equipment Disassembly	30 min.	26 min.	35 min.	30 min.
Coding tapes and preparing them for shipment	10 min.	9 min.	10 min.	10 min.

prior to taping, 3 minutes to change the tapes during the middle of the session being recorded, 30 minutes to disassemble the equipment after the taping was completed, and 10 minutes to code the tapes for the scorers. The total time required for the taping of a given session counting all operations plus the actual meeting time was approximately 177 minutes. A similar analysis was made of the tasks required in scoring the tapes once they were produced. The results of this evaluation can be found in Table XI. As noted in this table the average times for the operations required in scoring the tapes were given as follows: 12.75 minutes to thread the tape and prepare log and score sheet, 58 minutes to listen to the complete tape and rewind prior to scoring, 86 minutes to score the tape, 16 minutes to rewind the tape and total up the frequencies of response per-category, 39 minutes to total frequencies for all subjects and enter data on summary sheets, and 9 minutes to fill out log. The total average time required to score one tape was approximately 260 minutes.

The time required to administer a number of other measures to be used in the subsequent phases of the research were also

TABLE XI

Time Motion Analysis of Tasks
Required for Videotape Scoring

	Scorer								Average Time
	1		2		3		4		
Session	1	2	1	2	1	2	1	2	
Thread tape; prepare log & score sheet	12	10	14	15	15	12	10	14	13 mins.
Listen to tape; rewind.	62	55	51	60	56	60	60	60	58 mins.
Score tape.	80	75	90	90	95	85	90	85	86 mins.
Rewind tape, count frequencies.	17	15	18	12	16	18	17	15	16 mins.
Total frequencies; enter on summary sheet in duplicate.	45	40	40	35	35	40	42	40	40 mins.
Enter data in log.	10	8	8	10	12	10	10	10	10 mins.
Discuss criteria.	--	90	--	60	--	90	--	60	75 mins.
Total	226	293	221	282	229	315	229	284	260 mins.

evaluated. The results of this aspect of the feasibility study can be summarized in the following manner: 30 minutes, approximate time required to administer the short form of the WAIS; 45 minutes, average time required for the administration of the Spilka Alienation Measure; 15 minutes, average time required to fill out the Revised Behavior Inventory. The last of these measures must be filled out by someone other than the subject himself and will be used with treatment facility staff, school personnel, and employers.

DISCUSSION

Prior to discussing the implications of the specific results presented in the previous section, several general comments should be made about the nature of the research effort carried out during the first phase of operations. Since one of the main purposes of the Phase I activities was to experiment with the procedures designed to assess the underlying processes taking place during GGI, a great deal of modification was carried out with respect to the basic analysis system through out the period when data collection and scoring were taking place. In all cases where such modifications were made, they were made in the

interest of refining the procedures in order to derive a more accurate representation of the dynamics taking place in the GGI groups. All of the modifications stemmed from problems encountered in the use of the system, and many were nothing more than the clarification of the meaning of a specific definition for one of the scoring categories. Because of the degree of importance placed on the Revised Behavior Scores System relative to the planned assessment of the effectiveness of GGI techniques in rehabilitating delinquents, and because of the lack of information about the efficacy of this measure, the majority of emphasis during the first phase was placed on evaluating the utility of this instrument rather than on the other measures to be used in the subsequent phases of the research program. Some justification for this approach came from the fact that the other measures had been used in a number of previous research endeavors and their characteristics had already been described.

With respect to the specific findings relative to the assessment of the reliability for the six basic response mode categories the following comments seem appropriate. In general, the data generated across scorers for these classifications of

behavior can be considered highly sound. On several occasions the data presented in Tables I through III showed some variation with respect to reliabilities for a given response category from session-to-session. A number of these differences can be explained in terms of the constant modification taking place throughout the data collection period. The two classes of behavior that presented the most difficulty were Categories V: Supportive Acts, and VI: Assertive Supportive Acts. The problem involved in using these two categories stemmed from the fact that the definitions of the behaviors to be included in these classes were very similar. Since both deal with the agreement of one individual in the group with the position taken by another member, some confusion developed over the appropriate category to be used for a number of the behaviors viewed on the tapes. This problem was resolved by including only those statements that expressed general support without the use of a strong personal referent in Category V: Supportive Acts. In contrast, only those statements containing or implying a strong personal referent were scored in Category VI: Assertive Supportive Acts. As noted in the previous section the two categories with the highest

and most stable reliabilities were Categories I: Neutral Communication, and II: Assertive or Dominant Acts. One of the primary factors underlying these findings was that these two scoring classes contained over 50 percent of the total responses made by all group members. The category with the least number of responses consistently was Category IV: Withdrawal Acts. One possible explanation for the lack of these types of behavior, centers on the fact that the group will not tolerate these kinds of responses within the GGI sessions. If withdrawal from the program takes place it probably takes place outside of the confines of the actual treatment meetings. One hypothesis that can be formulated on the basis of this assumption could be that the group member who finds his withdrawal behavior being suppressed in the GGI groups has a higher probability of running away from the treatment facility and/or becoming an isolate with respect to the activities taking place at the treatment setting.

Some of the most interesting results found in the present study relate to the attempt to establish the reliabilities for the use of the reinforcement contingencies in relation to the responses scored for a specific category. With respect to the

positive reinforcement contingency a number of comments can be made. First, it was found that only the first two categories received sufficient numbers of responses that were positively reinforced to justify computing the reliability coefficients. This finding was not totally surprising when the content of the other factors was considered. Categories III: Antagonist Acts, and IV: Withdrawal Acts both contain predominantly those responses that the GGI groups consider to be of negative value. Therefore, it would be expected that these behaviors would receive little support in terms of positive reinforcement when they do occur in the interaction taking place in GGI. With respect to the findings for Categories V: Supportive Acts, and VI: Assertive Supportive Acts, another explanation was needed in that the responses scored in these categories generally reflect behavior that the groups positively value. The lack of positive reinforcement given to responses scored in these two categories stemmed from the fact that these behaviors were supportive of another group members position. After the type V, or VI response was made the individual being supported generally would regain direction of the group and continue the position he was

advocating prior to receiving the support, without directly acknowledging the supportive behavior. Therefore, few of these responses received reinforcement, even though they themselves were used as positive reinforcers. In fact the responses scored in these two categories represented the total positive reinforcement provided in the GGI groups.

Several additional comments must be made relative to the reliability data generated for the positive reinforcement contingency. The variability noted in Tables IV, V, and VI with respect to differences between the coefficients for the first session and those for the remaining sessions can be explained in terms of the lack of sufficient training on the part of the scorers at this point in data collection and analysis. Because of the timing of the initiation of the project, the scorers still had not achieved adequate proficiency with respect to using the criteria for the reinforcement contingency components of the scoring system. However, it should be noted that by the last session the reliabilities for the first two response categories were well within the acceptable limits for data interpretation. One of the most interesting findings of the present

research relates to the tape scored for Session VI with respect to Category II: Assertive or Dominant Acts. This was the only occasion in which the responses scored for this category were not followed by even one positive reinforcement. Therefore, there was perfect agreement between the scorers due to the lack of reinforcements.

The data for the negative reinforcement contingency were in many ways quite similar to those discussed for the positive reinforcements. Again the most stable patterns of reliability were for the first two categories of response. However, in the case of the negative reinforcement Category III: Antagonistic Acts, also received sufficient reinforcement to justify the computation of reliability coefficients. Even though the first two categories of response were primarily composed of those types of behavior that frequently would be considered positive or at least neutral, they still received a large amount of negative reinforcement. In fact the use of the negative reinforcement far exceeded the use of positive reinforcement to the extent of a five-to-one ratio for several of the sessions analyzed. Part of the explanation for this finding probably rests on

the fact that the members of the group felt that a given individual was using the behaviors scored in Categories I, and II as a means of manipulating the group through conning. Another aspect related to this finding was that the attitude of the group relative to appropriate behavior. In some cases it appeared as if the group felt that if an individual was doing something wrong he must be corrected immediately. With respect to the third class of responses, Category III: Antagonistic Acts, a different explanation seems appropriate. These behaviors in the main were those that the group put the highest negative value on and consequently made a concerted attempt to suppress them through the use of negative reinforcement. As was noted earlier these behaviors, unlike those scored in the first two categories, rarely receive positive reinforcement from the group members. Some further comment must also be given to the fourth class of behavior, Category IV: Withdrawal Acts, in that by the last session the scorers were reliably assessing this category in terms of negative reinforcement. This was due to a clarification of the criteria of this reinforcement contingency in the later stages of the data collection period to

include those behaviors that were punishing as well as negatively reinforcing in this category.

With respect to the consistency of the reliability data for the negative reinforcement contingency a number of comments can be made. First, there were large differences in the magnitude of the reliability coefficients between the first session and those that followed. These differences can be explained using the same rationale presented for a similar finding for the positive reinforcement data, that in effect the scorers were not thoroughly profficient in using the criteria for scoring the negative reinforcements by the start of the data collection and analysis period. Next, the use of the negative reinforcement contingency with the fourth category presented several interesting problems. The responses scored in Category IV: Withdrawal Acts would be expected to elicit negative reinforcement on the part of other members, however, due to the lack of sufficient numbers of responses in this category it was impossible to accurately assess the reliability of scoring negative reinforcement. In the one case where enought responses were scored in this category, Session VII, the reliability for the

with respect to the qualitative interpretation of the data when they were used with a given response. This subjective finding seems quite appropriate in that a broad range of different behaviors can be included in any one of the response mode categories. However, it should be emphasized that the use of the surscores and scoring convention should be limited to a descriptive function, and not be used as quantifiable data. The most widely used surscore by the scorers for the GGI data was the Tension Surscore, with the Task Determining Acts and Group Maintaining Acts Surscores next in order of usage. The least used of the surscores was the one dealing with unpredictable behavior which was rarely used with the present set of data. The scoring convention for false starts was used a great deal during the initial sessions, but diminished to some extent after the scorers became increasingly familiar with the language patterns of the subjects. In that the data included responses made by southern rural subjects, both white and black, some adjustment was necessary with respect to the listening habits of the scorers.

Since one of the goals of the present research was to assess whether or not it was even possible to carry out the tasks

required to evaluate therapy programs in terms of the processes taking place in therapy, the changes resulting from such processes, and the durability of such changes, a set of time motion studies was initiated with respect to the various components of the project design. In terms of the data presented in Tables X and XI, the implementation of the basic tasks required to use the Revised Behavior Scores System can be considered quite feasible. In fact, when the project was first begun it was felt that the scorers would have to score the tape for the responses made by three or at the most four of the subjects in a group, and then replay it for another set of subjects until all of the members of a given group were evaluated. However, after the training period was completed, all of the scorers were able to analyze the behavior of all of the members of a given group simultaneously. With respect to the videotape production operation several important factors must be elaborated on. First, it was felt that the specialized training that the videotape assistant received was of great importance in minimizing the amount of time wasted as a result of equipment breakdown. Normally, equipment such as that used in the present study requires

a great deal of attention with respect to maintenance outside of the settings where it must be used. However, in the present case there was not one instance when the equipment had to be removed from the treatment location because of failure. The results of the time-motion study for this aspect of the research clearly demonstrated that one individual could adequately handle all of the tasks related to videotape production.

The information contained in the videotape assistant's and scorers' logs provided a number of interesting observations that should be considered. Both the scorers and the tape assistant noted the importance of having meetings with each other in order to share ideas relative to the best possible camera placements, microphone settings, etc. On several occasions the quality of the production of the tapes was markedly increased after such meetings took place. During these conferences the individuals responsible for taping and scoring reviewed tapes that had already been used with the scorers pointing out their needs to the tape technician. Another important result of the information recorded in the logs was the scorers' general finding that it was necessary to have a substantial break between scoring two tapes because of the fatigue effects related to the

task. Preferably no more than two tapes or one complete session should be scored in a single day. Since the scorers were responsible for other tasks related to the research project, there was no great loss in efficiency in terms of deriving maximum use of personnel while still meeting the above requirement of scoring only two tapes per-analysis session. Because of comments made by the scorers in the logs, the videotape equipment used for scoring was altered in line with the auditory qualities of the tape. It was found that it was advisable to use a set of earphones equipped with a device that enabled the scorers to vary the sound level of the tapes they were scoring. In those instances when the entire group being taped responded to an individual in the GGI setting, the high impedance microphone used in taping picked up all of the sound with the result being that it was quite difficult to isolate who was making what response. By using the specially equipped earphones, it was possible to lower the noise level to a point where such discriminations could be made by the scorers.

In general, the results of the present study are felt to provide evidence that data can be collected that reflect the

basic processes taking place within the GGI sessions that are responsible for behavior change. The analysis system applied to the group settings in the present study, derived from the work of individuals in the areas of group dynamics and learning, shows considerable promise for isolating the basic variables that must be considered if group therapy techniques and approaches are to be adequately evaluated. The results discussed above indicate that the system should yield highly reliable data relative to the behavior occurring in therapy groups while at the same time providing measures of the learning processes accounting for such behavior. Therefore, one of the important implications of the present study is that the findings of basic laboratory research in the area of learning can be applied in a meaningful manner to evaluate treatment in a field setting where the behavior changes occurring frequently are brought about without any consideration to basic learning principles.

SUMMARY

The main findings of the present study relate to the establishment of the reliabilities for the main components of the Revised Behavior Scores System developed to assess the basic processes taking place in settings using Guided Group Interaction as a theraputic technique. This system is derived from the scoring categories, surscores, and scoring conventions developed by Borgatta and Crowther (1965) in their Behavior Scores System. The procedures discussed by these authors are modified to contain components for the measurement of positive and negative reinforcement contingencies. These modifications are made in an attempt to assess the basic learning processes occurring in Guided Group Interaction that account for subsequent behavior change. The results indicate that in general the system is sufficiently reliable to generate sound data with respect to the six basic response categories (I: Neutral Communication, II: Assertive or Dominant Acts, III: Antagonistic Acts, IV: Withdrawal Acts, V: Supportive Acts, and VI: Assertive Supportive Acts). In addition the data for the reinforcement contingencies indicates that these components of the system

are reliable if used with the appropriate classes of behavior. Positive reinforcement should be used with Categories I: Neutral Communication, and II: Assertive or Dominant Acts. In that responses for the other categories are seldom reinforced in a positive manner, and that those that do receive positive reinforcement are generally scored in the first two categories, it is felt that this component should provide a great deal of important information relative to the dynamics taking place in treatment. Similarly, the reliabilities for the negative reinforcement are sufficiently high with respect to the first three response categories to yield data of some utility in evaluating group therapy programs.

The data generated for the surscores and the one scoring convention used in the Revised Behavior Scores System indicates that these components in the analysis system should not be used for gathering quantifiable data. Their main function is descriptive with respect to defining the behavior taking place in GGI as scored by the research assistants. The reason for this conclusion is that these indices do not occur frequently enough to evaluate their reliability and therefore their usage may be

suspect. Since one of the main criteria used in evaluating research is the reliability of the measures used it is advisable not to use these aspects of the Revised Behavior Scores System in trying to quantify the behavior taking place in a group therapy setting.

In addition to the reliability data, information was gathered with respect to the amount of time and effort required to implement the procedures for using the analysis system commented on above. It was found that the tasks required for the gathering and analysis of data were quite feasible and did not provide any great difficulty for the project staff. This aspect of the research yielded much valuable information relative to refining the specific tasks used in data collection, preparation and analysis.

REFERENCES

Borgatta, E. F. & Crowther, B. A workbook for the study of social interaction processes. Chicago: Rand McNally & Co. 1965.

Ferster, C. B. The use of the free operant in the analysis of behavior. Psychological Bulletin, 50: 263-275, 1953.

Ferster, C. B. & Skinner, B. F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.

Glasser, Wm. Reality therapy. New York: Harper & Row, 1965.

Jacobs, M. & Christ, J. Structuring and limit-setting as techniques in the group treatment of adolescent delinquents. Community Mental Health Journal, 1967, 3(3), 237-244.

Knight, D. The Marshall Program: Assessment of a short-term institutional treatment program. Part I: Parole outcome and background characteristics. California Youth Authority: Research Report, (March), 1969, No. 56.

Knight, D. The Marshall Program: Assessment of a short-term institutional treatment program. Part II: Amenability to confrontive peer group treatment. California Youth Authority: Research Report, (August), 1970, No. 59.

Peins, M. Client-centered communication therapy for mentally retarded delinquents. Journal of Speech and Hearing Disorders, 1967, (32) 2, 154-161.

Persons, R. W. Psychological and behavioral change in delinquents following psychotherapy. Journal of Clinical Psychology, 1966, 22, 337-340.

Peterson, D. R., Quay, H. C. and Cameron, G. R. Personality and background factors in juvenile delinquency as inferred from questionnaire responses. Journal of Consulting Psychology, 1959, 23, 395-399.

Sandhu, H. S. Group sessions in a reformatory school in the Punjab (India). Corrective Psychiatry and Journal of Social Therapy, 1966 12 (5), 393-403.

Shore, M., Massimo, J. and Mack, R. Changes in the perception of interpersonal relationships in successfully treated adolescents. Journal of Consulting Psychology, 1965, (27) 1, 213-217.

Sidman, M. Tactics of scientific research: Evaluating experimental data in psychology. New York: Basic Books. 1960.

Skinner, B. F. Some contributions of an experimental analysis of behavior to psychology as a whole. American Psychologist, 8: 69-78, 1953.

Skinner, B. F. The experimental analysis of behavior. American Scientist, 45: 343-371, 1957.

Spilka, B. Alienation and Achievement among Ogalala Sioux Secondary School Students. Technical Report, August, 1970, Contract Number MH 11232, National Institute of Mental Health.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill Book Co., 1962.